

# An Analysis of How Spatiotemporal Dynamic Models of Brain Activity Could Improve MEG/EEG Inverse Solutions

Camilo Lamus, Matti S. Hämäläinen, Emery N. Brown, *Fellow, IEEE* and Patrick L. Purdon,

**Abstract**—MEG and EEG are noninvasive functional neuroimaging techniques that provide recordings of brain activity with high temporal resolution, and thus provide a unique window to study fast time-scale neural dynamics in humans. However, the accuracy of brain activity estimates resulting from these data is limited mainly because 1) the number of sensors is much smaller than the number of sources, and 2) the low sensitivity of the recording device to deep or radially oriented sources. These factors limit the number of sources that can be recovered and bias estimates to superficial cortical areas, resulting in the need to include a priori information about the source activity. The question of how to specify this information and how it might lead to improved solutions remains a critical open problem. In this paper we show that the incorporation of knowledge about the brains underlying connectivity and spatiotemporal dynamics could dramatically improve inverse solutions. To do this, we develop the concept of the *dynamic lead field mapping*, which expresses how information about source activity at a given time is mapped not only to the immediate measurement, but to a time series of measurements. With this mapping we show that the number of source parameters that can be recovered could increase by up to a factor of  $\sim 20$ , and that such improvement is primarily represented by deep cortical areas. Our result implies that future developments in MEG/EEG analysis that model spatiotemporal dynamics have the potential to dramatically increase source resolution.

**Index Terms**—Inverse problem, Magnetoencephalography/Electroencephalography, Spatiotemporal statistics.

## I. INTRODUCTION

MAGNETOENCEPHALOGRAPHY (MEG) and electroencephalography (EEG) are functional neuroimaging tools that provide noninvasive recordings of the magnetic and electric fields at the scalp generated by neuronal currents. MEG and EEG hold particular promise as tools

to noninvasively study fast time-scale neural dynamics of human brain function in health and disease [1]. This makes MEG/EEG unique amongst other functional neuroimaging techniques such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), or diffuse optical tomography (DOT), which instead provide indirect measures of brain activity related to slower neurovascular changes. However, to properly interpret MEG and EEG recordings and fully realize their potential, one needs to estimate the source current activity underlying the measured electric potentials and magnetic fields at the scalp surface, i.e., to solve the ill-posed MEG/EEG inverse problem. Although significant progress has been made on this problem in the past few decades [2]–[13], at present MEG/EEG inverse solutions are rough estimates with poor spatial resolution (in the order of a few cm), that are insensitive to the majority of deep cortical and subcortical regions.

The accuracy of source estimates is limited in part due to two main factors: the dimensionality and the biophysics in MEG/EEG inverse problem. These factors can be understood by analyzing the *lead field matrix*  $\mathbf{X}$ , which maps the cortical activity of a few thousand dipole sources,  $\beta_t$ , to the recordings in a few hundred scalp sensors,  $\mathbf{y}_t$ , at an individual time instant  $t$  [1], [14]:

$$\mathbf{y}_t = \mathbf{X}\beta_t + \text{"noise"}.$$

In the case of the dimensionality, since the number of sources to be estimated is an order of magnitude larger than the number of sensors, different source configurations can produce identical scalp recordings, making solutions to this inverse problem non-unique. As a result, the number of sources that we can expect to recover from measurements in an individual time point  $t$ , which is determined by the rank of  $\mathbf{X}$ , is much smaller than the number of sources. The second factor is due to the biophysics of MEG/EEG. The amplitude of the scalp fields and potentials rapidly decays with the inverse square of the distance from the sensor to the source [1]. Therefore, the sensitivity of the recording device—the signal power generated by an individual active source across all sensors given by the norm of the columns of  $\mathbf{X}$ —is very low for a large percentage of cortical areas. The geometry of the electromagnetic fields generated by current dipoles poses additional challenges. The measured fields tend to be widely distributed across the scalp, and radially-oriented sources can be magnetically silent [15]. Overall, the dimensionality and biophysics underlying MEG

C. Lamus is with the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: camilo@neurostat.mit.edu).

M.S. Hämäläinen is with Harvard Medical School, Boston, MA 02115 USA, with the Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA 02129 USA, and with the Department of Neuroscience and Biomedical Engineering, Aalto University School of Science, Espoo, Finland.

E.N. Brown is with the Department of Brain and Cognitive Sciences and the Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA, with the Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, MA 02114 USA, and with Harvard Medical School, Boston, MA 02115 USA.

P.L. Purdon is with Harvard Medical School, Boston, MA 02115 USA, and with the Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, MA 02114 USA (e-mail: patrickp@nmr.mgh.harvard.edu)

Manuscript received November 11, 2015; revised Month Day, Year.

and EEG limit the number of source parameters that can be recovered, and effectively restrict estimates to cortical areas whose activity is most easily detected by the sensors, ultimately imposing the need to include a priori information about the source currents. The question of how to specify this a priori information and how it might improve solutions is one of the most fundamental problems in bioelectromagnetism and neuroimaging.

Prior information about the current dipole distribution has been used in the MEG/EEG inverse problem literature in order to obtain unique estimates. This information has taken the form of a probabilistic model or optimization penalty that implicitly or explicitly assumes cortical activity is either independent across time [2]–[5], temporally or spatio-temporally smooth [6]–[10], or follow a linear dynamic process [11]–[13]. While these priors alleviate issues related to the non-uniqueness of source estimates, they do not necessarily improve on the limitations that stem from the rank deficiency and restricted sensitivity of the lead field matrix. However, one way to ameliorate these issues could be achieved by incorporating prior knowledge about the spatiotemporal physiological relations present in brain activity [16]–[19]. This is because in such a spatiotemporal system, information about source activation at a given time instant  $t$  is contained not only in the immediate measurement  $\mathbf{y}_t$  but also in multiple observations over a time interval  $[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$ . As a result of this dynamic flow of information, we can imagine a mapping that relates brain activity at a given moment in time to the measurements in the complete analysis interval—a *dynamic lead field mapping*. The rank and sensitivity of the dynamic lead field mapping could be substantially better than those of the static lead field matrix  $\mathbf{X}$ , not only because of its increased dimensionality, but also because of how information can flow from brain areas that are harder to detect at the scalp to those that are easier to detect.

Inspired by electrophysiology and neuroanatomy studies [16]–[19], in this paper we show how the incorporation of information about the source spatial connectivity and dynamics could dramatically increase both the number of sources that can be recovered, as well as the sensitivity for detecting such sources. To do this, we develop the concept of the *dynamic lead field mapping*, a dynamic extension of the lead field matrix that allows us to analyze the rank and sensitivity of the mapping between source activity at a given time instant to the complete set of measurements both forward and backwards in time. This dynamic lead field mapping, though developed for the specific problem of MEG/EEG source imaging, is firmly grounded in dynamic systems theory. With this mapping we show that the number of sources that can be effectively recovered increases by up to a factor of  $\sim 20$  by modeling the most basic local cortical dynamic connections. Furthermore, we show that the inclusion of such local cortical connections increases the sensitivity for detecting sources distributed across the brain, and that the increase in sensitivity is more pronounced in sources located within sulci and other deeper areas. At the core of our technical development is a projection operation that allows us to analyze the rank and sensitivity resulting from different prior source models. In particular, we

analyze a space-time separable model and a static model that assumes temporal independence, both of which are frequently used in MEG/EEG analysis. We find that improvements in the number and sensitivity of sources that can be recovered occur only when dynamic spatiotemporal connections are modeled. Our results imply that future developments in MEG/EEG analysis that explicitly model dynamic connections between brain areas have the potential to dramatically increase spatiotemporal resolution by taking full advantage of the brain's inherent spatiotemporal structure, connectivity, and dynamics.

## II. METHODS

### A. The spatiotemporal dynamic source model and the MEG/EEG measurement model

Converging lines of evidence suggest that brain activity is a spatiotemporal dynamic process that is organized in part by structural connections at different scales. At the smaller spatial scale, intracranial recordings in different species have revealed that cortical activity exhibits strong correlations that persist up to a distance of 10 mm during rest and task periods [16]–[18]. These local cortico-cortical dynamic interactions can be supported neuroanatomically by the fact that axonal collateral projections from pyramidal cells spread laterally approximately 6 mm along the cortical surface [19]. In addition, long-distance correlations can exist by means of white matter tracts that connect distant brain regions, like those which are thought to support large-scale brain network activity such as the resting-state networks studied via fMRI [20]–[22].

In order to provide a conservative analysis where spatiotemporal dynamics consistent with neurophysiological evidence are considered, we choose to focus on a parsimonious model that incorporates local spatiotemporal interactions. One way to model local spatiotemporal connections of this type is to use a first order linear dynamic process. In this model, neuronal currents at a given point in time  $t$  and spatial location  $i$ ,  $\beta_{i,t}$ , are a function of past neuronal currents at the same location,  $\beta_{i,t-1}$ , as well as past currents,  $\beta_{j,t-1}$ , at locations  $j$  within a local neighborhood  $\mathcal{N}(i)$ :

$$\beta_{i,t} = f_{i,i}\beta_{i,t-1} + \sum_{j \in \mathcal{N}(i)} f_{i,j}\beta_{j,t-1} + \omega_{i,t}. \quad (1)$$

In Equation (1), the weights  $f_{i,j}$  represent the interaction between sources at locations  $i$  and  $j$ . At each location  $i$ , the weights  $f_{i,j}$  corresponding to its neighbors  $j \in \mathcal{N}(i)$  are assumed to be positive and inversely proportional to the distance between sources. In addition, they are normalized such that the contribution of the neighbors to the dynamics of the  $i$ th source equals its self contribution, while the total contribution is constant  $0 < \phi < 1$  in order to obtain stable source dynamics, i.e.,  $\sum_{j \in \mathcal{N}(i)} f_{i,j} = f_{i,i}$  and  $\sum_{j \in \mathcal{N}(i)} f_{i,j} + f_{i,i} = \phi < 1$ . Furthermore, the input process  $\omega_{i,t}$  is assumed to be Gaussian with zero-mean and independent across time and space. The spatiotemporal model in Equation (1), which has been previously used in dynamic source localization analysis [11], [13], can be readily expressed in vector form as:

$$\boldsymbol{\beta}_t = \mathbf{F}\boldsymbol{\beta}_{t-1} + \boldsymbol{\omega}_t, \quad (2)$$

where  $\beta_t$  is a vector of dimension  $p$  ( $\sim 10^3$ ), and the input process  $\omega_t$  is Gaussian with zero mean and independent across time with spatial covariance matrix  $\mathbf{Q}$ , which we assume to be diagonal and positive definite.

In an MEG/EEG experiment, we obtain a recording of the magnetic field and electric potential from hundreds of sensors located on or above the scalp at times  $t \in [1, 2, \dots, T]$ . At time  $t$ , the resulting data vector  $\mathbf{y}_t$ , which is of dimension  $n$  ( $\sim 10^2$ ), is related to the source vector  $\beta_t$  by the observation equation [1], [14]:

$$\mathbf{y}_t = \underbrace{\mathbf{X}\beta_t}_{\text{Signal}} + \underbrace{\varepsilon_t}_{\text{Noise}}, \quad (3)$$

where  $\mathbf{X}$  is the  $n \times p$  lead field matrix computed using a quasistatic approximation of the Maxwell's equations, and  $\varepsilon_t$  is the Gaussian white noise vector with zero mean and spatial covariance  $\mathbf{R}$  representing background instrumental and environmental noise. In Equation (3),  $\mathbf{X}\beta_t$  represents the *signal* portion of the model, and the *noise* term  $\varepsilon_t$  is independent from  $\beta_t$  for all time points.

### B. The dynamic lead field mapping

In the observation model (Eq. 3), the signals generated by brain activity are represented by the product of the lead field matrix and the source vector:  $\mathbf{X}\beta_t$ . Since the resulting product and the source vector are independent of the noise term  $\varepsilon_t$ , the lead field matrix  $\mathbf{X}$  contains all the information related to the mapping of the brain source vector  $\beta_t$  at a particular point in time  $t$  to the measurement  $\mathbf{y}_t$  at that same point in time  $t$  (Figure 1A). From this static point of view, it is clear that the maximum number of independent variables that could be determined from a single measurement in time is limited by the rank of the  $\mathbf{X}$  matrix, which unfortunately is less than or equal to the number of sensors, and much smaller than the number of sources [23]. However, if we consider the brain's source activity as a spatiotemporal dynamic process, we can dramatically improve on these limitations.

In a spatiotemporal system, the information about the source vector  $\beta_t$  at a particular time  $t$  is contained not only in the immediate observation  $\mathbf{y}_t$  but also in the previous  $[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{t-1}]$  and future  $[\mathbf{y}_{t+1}, \mathbf{y}_{t+2}, \dots, \mathbf{y}_T]$  observations. Because of this, information from the source vector  $\beta_t$  is effectively mapped to the complete time series of measurements through a function (Figure 1B), which we call the *dynamic lead field mapping*  $\mathbf{D}_t$ , whose rank and sensitivity would be greater than those of the static lead field matrix  $\mathbf{X}$ . As we will show below, this mapping can be expressed as:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_T \end{bmatrix} = \underbrace{\mathbf{D}_t \beta_t}_{\text{Signal}} + \underbrace{\mathbf{n}_t}_{\text{Noise}}, \quad (4)$$

where the noise term  $\mathbf{n}_t$  has dimension  $nT$ . In analogy to the lead field matrix  $\mathbf{X}$  in the *static* observation model (Eq. 3), the *dynamic* lead field mapping  $\mathbf{D}_t$  can be used to determine the number of independent variables we can recover as well

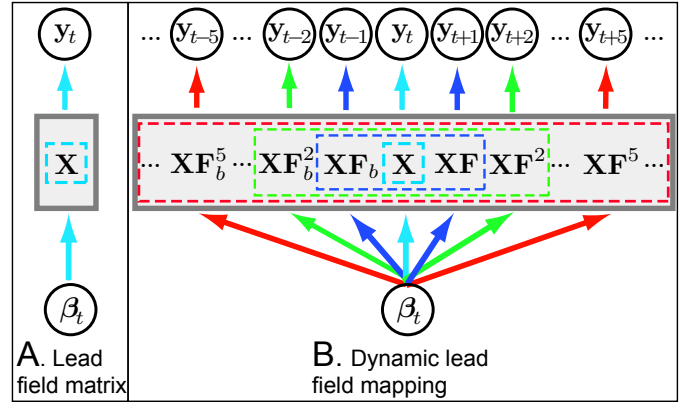


Fig. 1. **The dynamic lead field mapping.** A: The (static) lead field matrix  $\mathbf{X}$  determines how information from the source vector  $\beta_t$  propagates to the measurement  $\mathbf{y}_t$  at the same point in time  $t$ . B: When spatiotemporal dynamics are modeled, the dynamic lead field mapping  $\mathbf{D}_t$  determines how information from the source vector  $\beta_t$  at a given time  $t$  is mapped to the complete time series of measurements  $[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$ .

as the sensitivity for detecting such sources when we consider the complete time series of observations in a spatiotemporal dynamic framework.

To obtain the dynamic lead field mapping  $\mathbf{D}_t$  in Equation (4), we must derive an observation model for the complete set of measurements  $[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$  where: i) the *signal* portion of the model is a function of only the source vector  $\beta_t$  at a particular time  $t$ , and ii) the *noise* term is independent of the source vector  $\beta_t$ . We will consider three separate cases depending on whether the observation vector  $\mathbf{y}_{t \pm k}$  ( $k = 0, 1, \dots$ ) corresponds to a future, present, or past observation with respect to the source vector  $\beta_t$  at the present time  $t$ . Because of this partition we will consider the dynamic lead field mapping  $\mathbf{D}_t$  as a block matrix composed of submatrices  $\mathbf{P}_{t \pm k, t}$ :

$$\mathbf{D}_t = \begin{bmatrix} \mathbf{P}_{1,t} \\ \mathbf{P}_{2,t} \\ \vdots \\ \mathbf{P}_{T,t} \end{bmatrix}.$$

We call the blocks  $\mathbf{P}_{t \pm k, t}$  the projection matrices.

In the case where the measurement vector corresponds to the present observation ( $k = 0$ ), we use the static observation model (Eq. 3) to obtain the relation between  $\beta_t$  and  $\mathbf{y}_t$  in the block number  $t$  of Equation (4):

$$\mathbf{y}_t = \underbrace{\mathbf{X}\beta_t}_{\text{Signal}} + \underbrace{\varepsilon_t}_{\text{Noise}}.$$

From the equation above we can identify that  $\mathbf{P}_{t,t} = \mathbf{X}$ .

To obtain the projection matrices corresponding to future observations, we start by using the recursion defining the source vector spatiotemporal dynamics (Eq. 2) and iterate it  $k \in [1, 2, \dots]$  times to obtain:

$$\begin{aligned}
\beta_{t+1} &= \mathbf{F}\beta_t + \omega_{t+1} \\
\beta_{t+2} &= \mathbf{F}^2\beta_t + \mathbf{F}\omega_{t+1} + \omega_{t+2} \\
&\vdots \\
\beta_{t+k} &= \mathbf{F}^k\beta_t + \sum_{j=1}^k \mathbf{F}^{k-j}\omega_{t+j}.
\end{aligned} \tag{5}$$

Using this  $k$ -step iteration into the future along with the static measurement model (Eq. 3), we obtain the relation between  $\beta_t$  and the future measurement  $\mathbf{y}_{t+k}$ :

$$\mathbf{y}_{t+k} = \underbrace{\mathbf{X}\mathbf{F}^k\beta_t}_{\text{Signal}} + \underbrace{\sum_{j=1}^k \mathbf{X}\mathbf{F}^{k-j}\omega_{t+j}}_{\text{Noise}} + \varepsilon_{t+k}. \tag{6}$$

We should note that Equation (6) represents the  $(t+k)$ th block of Equation (4), and therefore we identify that  $\mathbf{P}_{t+k,t} = \mathbf{X}\mathbf{F}^k$ .

To derive the past projection matrices we make use of an equivalent representation of the source dynamics (Eq. 2) known as the *backwards Markovian* model [24]:

$$\beta_t = \mathbf{F}_b\beta_{t+1} + \omega_t^b. \tag{7}$$

In this representation, where the state vector  $\beta_t$  evolves backwards in time, the time-reversed transition matrix is given by  $\mathbf{F}_b = \mathbf{C}\mathbf{F}'\mathbf{C}^{-1}$ , where  $\mathbf{C} = \text{Cov}(\beta_t)$  is the steady-state source covariance (see Section V-A in **Supplemental Materials** for details). The backwards input process  $\omega_t^b$ , which is independent across time and Gaussian, has zero mean and covariance matrix  $\mathbf{Q}_b = (\mathbf{C} - \mathbf{F}_b\mathbf{C}\mathbf{F}_b')$ . With this equivalent backwards dynamic representation (Eq. 7), we proceed to obtain the relation between  $\beta_t$  and the previous measurement  $\mathbf{y}_{t-k}$  just as we did when we considered future measurement. In this case the  $k \in [1, 2, \dots]$  step backwards iteration is given by:

$$\begin{aligned}
\beta_{t-1} &= \mathbf{F}_b\beta_t + \omega_{t-1}^b \\
\beta_{t-2} &= \mathbf{F}_b^2\beta_t + \mathbf{F}_b\omega_{t-1}^b + \omega_{t-2}^b \\
&\vdots \\
\beta_{t-k} &= \mathbf{F}_b^k\beta_t + \sum_{j=1}^k \mathbf{F}_b^{k-j}\omega_{t-j}^b.
\end{aligned} \tag{8}$$

We use the static observation model (Eq. 3) along with the  $k$ -step past iteration to obtain the  $(t-k)$ th block of Equation (4):

$$\mathbf{y}_{t-k} = \underbrace{\mathbf{X}\mathbf{F}_b^k\beta_t}_{\text{Signal}} + \underbrace{\sum_{j=1}^k \mathbf{X}\mathbf{F}_b^{k-j}\omega_{t-j}^b}_{\text{Noise}} + \varepsilon_{t-k}, \tag{9}$$

where we see that  $\mathbf{P}_{t-k,t} = \mathbf{X}\mathbf{F}_b^k$ .

At this point we are ready to explicitly write the complete observation model (Eq. 4) as well as the dynamic lead field mapping  $\mathbf{D}_t$ . For a particular time  $t$ , we use the equations mapping the source vector  $\beta_t$  to the present (Eq. 3), future

(Eq. 6), and past (Eq. 9) to obtain the complete observation model:

$$\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{t-1} \\ \mathbf{y}_t \\ \mathbf{y}_{t+1} \\ \vdots \\ \mathbf{y}_T \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{X}\mathbf{F}_b^{t-1} \\ \vdots \\ \mathbf{X}\mathbf{F}_b \\ \mathbf{X} \\ \mathbf{X}\mathbf{F} \\ \vdots \\ \mathbf{X}\mathbf{F}^{T-t} \end{bmatrix}}_{\mathbf{D}_t} \beta_t + \underbrace{\begin{bmatrix} \mathbf{e}_{1,t} \\ \vdots \\ \mathbf{e}_{t-1,t} \\ \mathbf{e}_{t,t} \\ \mathbf{e}_{t+1,t} \\ \vdots \\ \mathbf{e}_{T,t} \end{bmatrix}}_{\mathbf{n}_t}, \tag{10}$$

where error terms are given by,

$$\mathbf{e}_{t\pm k,t} = \begin{cases} \sum_{j=1}^k \mathbf{X}\mathbf{F}_b^{k-j}\omega_{t-j}^b + \varepsilon_{t-k}, & -k > 0 \\ \varepsilon_t, & k = 0 \\ \sum_{j=1}^k \mathbf{X}\mathbf{F}^{k-j}\omega_{t+j} + \varepsilon_{t+k}, & +k > 0 \end{cases}.$$

From the complete observation model (Eq. 10) we can see that the dynamic lead field mapping is:

$$\mathbf{D}_t = [\mathbf{F}_b^{t-1'}\mathbf{X}', \dots, \mathbf{F}_b'\mathbf{X}', \mathbf{X}', \mathbf{F}'\mathbf{X}', \dots, \mathbf{F}^{T-t'}\mathbf{X}']'. \tag{11}$$

In addition, we can observe that Equation (10) expresses the complete set of measurements  $[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$  as the sum of the *signal* function  $\mathbf{D}_t\beta_t$  and *noise* term  $\mathbf{n}_t = [\mathbf{e}'_{1,t}, \mathbf{e}'_{2,t}, \dots, \mathbf{e}'_{T,t}]'$ . We should note that the noise term  $\mathbf{n}_t$  is indeed independent of the source vector  $\beta_t$ , since the vectors  $\mathbf{e}_{t\pm k,t}$  correspond to the projection errors that result from projecting the measurements  $\mathbf{y}_{t\pm k}$  onto the source vector  $\beta_t$ . Such projections are obtained via the matrices  $\mathbf{P}_{t\pm k,t}$  (For details see Section V-B in **Supplemental Materials**). Consequently, the *dynamic* lead field mapping  $\mathbf{D}_t$  contains all the information describing how the brain source vector  $\beta_t$ , at a particular point in time  $t$ , propagates to the complete set of measurements. Therefore, we can use  $\mathbf{D}_t$  to determine the number of independent variables we can recover, as well as the sensitivity for detecting such sources, from the complete time series of observations in a spatiotemporal dynamic framework.

### C. Extension to general Gaussian source models

The expression for the dynamic lead field  $\mathbf{D}_t$  in Equation (11) stems from a specific model of spatiotemporal dynamics. In this section we derive a more general expression for  $\mathbf{D}_t$  for Gaussian spatiotemporal models. This will allow us to compare the properties of our dynamic source model to those of more commonly used models. In these more general models, the joint distribution of the sources has a zero mean and is Gaussian, as in our spatiotemporal dynamic source model (Eqs. 1 and 2), but the spatiotemporal covariance structure differs from that in our model. Specifically, under a model  $\mathbf{m}$ , the joint distribution of the source vectors  $[\beta_1^{(\mathbf{m})}, \beta_2^{(\mathbf{m})}, \dots, \beta_T^{(\mathbf{m})}]$  is Gaussian with zero mean ( $\mathbb{E}[\beta_t^{(\mathbf{m})}] = \mathbf{0}$ ) and arbitrary cross-covariances  $\mathbb{E}[\beta_k^{(\mathbf{m})}\beta_t^{(\mathbf{m})'}]$ , where  $\mathbb{E}$  denotes the expectation operator. The MEG/EEG measurement are obtained via the static observation model (Eq. 3):

$$\mathbf{y}_t^{(m)} = \mathbf{X}\beta_t^{(m)} + \varepsilon_t. \quad (12)$$

Just as we did in the case of the dynamic lead field mapping, when considering a model  $m$  we would like to express the complete time series of measurements  $[\mathbf{y}_1^{(m)}, \mathbf{y}_2^{(m)}, \dots, \mathbf{y}_T^{(m)}]$  as the sum of a function that only depends on the source vector  $\beta_t^{(m)}$  at the present time  $t$  and a noise term that is independent of  $\beta_t^{(m)}$ . As we will show below, this can be done via the linear relation,

$$\begin{bmatrix} \mathbf{y}_1^{(m)} \\ \mathbf{y}_2^{(m)} \\ \vdots \\ \mathbf{y}_T^{(m)} \end{bmatrix} = \underbrace{\mathbf{D}_t^{(m)} \beta_t^{(m)}}_{\text{Signal}} + \underbrace{\mathbf{n}_t^{(m)}}_{\text{Noise}}, \quad (13)$$

where  $\mathbf{D}_t^{(m)}$  is the equivalent of the dynamic lead field mapping for the model  $m$ , and the noise term  $\mathbf{n}_t^{(m)}$  is independent of the source vector  $\beta_t^{(m)}$ .

To obtain Equation (13) we again consider the matrix  $\mathbf{D}_t^{(m)}$  as composed of sub-matrix blocks  $\mathbf{P}_{t \pm k, t}^{(m)}$ :

$$\mathbf{D}_t^{(m)} = \begin{bmatrix} \mathbf{P}_{1, t}^{(m)} \\ \mathbf{P}_{2, t}^{(m)} \\ \vdots \\ \mathbf{P}_{T, t}^{(m)} \end{bmatrix}.$$

To obtain the  $t \pm k$ th block of Equation (13) we make use of the fact that, under a model  $m$ , the measurement vector  $\mathbf{y}_{t \pm k}^{(m)}$  at time  $t \pm k$  can be expressed as the sum of its projection onto the source vector  $\beta_t^{(m)}$  at time  $t$  and the projection error:

$$\mathbf{y}_{t \pm k}^{(m)} = \mathbf{P}_{t \pm k, t}^{(m)} \beta_t^{(m)} + \mathbf{e}_{t \pm k, t}^{(m)},$$

where the projection error  $\mathbf{e}_{t \pm k, t}^{(m)}$  is independent of the source vector at time  $t$ , and the projection matrix  $\mathbf{P}_{t \pm k, t}^{(m)}$  is the solution of the orthogonality equation<sup>1</sup>:

$$\mathbb{E} \left[ (\mathbf{y}_{t \pm k}^{(m)} - \mathbf{P}_{t \pm k, t}^{(m)} \beta_t^{(m)}) \beta_t^{(m)'} \right] = \mathbf{0}. \quad (14)$$

By making use of the static observation model (Eq. 12), it is easy to see that the projection matrices that solve the orthogonality equation (Eq. 14) under a given source model  $m$  are given by:

$$\mathbf{P}_{t \pm k, t}^{(m)} = \mathbf{X} \left( \mathbb{E}[\beta_{t \pm k}^{(m)} \beta_t^{(m)'}] \right) \left( \mathbb{E}[\beta_t^{(m)} \beta_t^{(m)'}] \right)^{-1}. \quad (15)$$

Therefore, for a given Gaussian model  $m$  with a specific spatiotemporal source covariance structure, we can compute the corresponding projection matrices  $\mathbf{P}_{t \pm k, t}^{(m)}$  and, just as we did in the development of the dynamic lead field mapping

<sup>1</sup>The projection error  $\mathbf{e}_{t \pm k, t}^{(m)} = \mathbf{y}_{t \pm k}^{(m)} - \mathbf{P}_{t \pm k, t}^{(m)} \beta_t^{(m)}$  is uncorrelated with  $\beta_t^{(m)}$ , namely  $\mathbb{E}[(\mathbf{y}_{t \pm k}^{(m)} - \mathbf{P}_{t \pm k, t}^{(m)} \beta_t^{(m)}) \beta_t^{(m)'}] = \mathbf{0}$  [24]. Since  $\mathbf{y}_{t \pm k}^{(m)}$  and  $\beta_t^{(m)}$  are jointly Gaussian, the projection error  $\mathbf{e}_{t \pm k, t}^{(m)}$  and  $\beta_t^{(m)}$  are also jointly Gaussian. Therefore, the projection error and the source vector are uncorrelated and jointly Gaussian, and thus are independent.

(Eqs. 4 and 10), obtain an equivalent observation model for the complete series of measurements by stacking up the projection matrices.

To give concrete examples of the  $\mathbf{D}_t^{(m)}$  matrix for general Gaussian source models, we will analyze two models commonly used in the source localization literature: i) the first, which we denote as the *IND* model, assumes the source vectors are independent across time [2], i.e.,  $\mathbb{E}[\beta_{t \pm k}^{(ind)} \beta_t^{(ind)'}] = \mathbf{0}$  for  $k \neq 0$ ; and ii) a space-time separable (*STS*) model in which the joint source covariance factors via the Kronecker product  $\mathbf{\Gamma} \otimes \mathbf{C}$  into a purely spatial covariance matrix  $\mathbf{C}$  and a purely temporal covariance matrix  $\mathbf{\Gamma}$  [7]. We should note that in the *STS* model, the covariance between source vectors at times  $t \pm k$  and  $t$  is given by  $\mathbb{E}[\beta_{t \pm k}^{(sts)} \beta_t^{(sts)'}] = \gamma_{t \pm k, t} \mathbf{C}$ , where  $\gamma_{k, t}$  is the element of the temporal covariance matrix  $\mathbf{\Gamma}$  at position  $t \pm k, t$ . With the covariance structure defined for these two models, and using Equation (15), we can see that the projection matrices for the *IND* and *STS* models are

$$\mathbf{P}_{t \pm k, t}^{(ind)} = \begin{cases} \mathbf{X} & k = 0 \\ \mathbf{0} & k \neq 0 \end{cases} \quad \text{and} \quad \mathbf{P}_{t \pm k, t}^{(sts)} = \frac{\gamma_{t \pm k, t}}{\gamma_{t, t}} \mathbf{X},$$

respectively. Therefore, the  $\mathbf{D}_t^{(m)}$  matrix for the *IND* and *STS* models, which constitute the equivalent of the  $\mathbf{D}_t$  matrix for these Gaussian models, are given by:

$$\mathbf{D}_t^{(ind)} = \mathbf{1}_t \otimes \mathbf{X} \quad \text{and} \quad \mathbf{D}_t^{(sts)} = \frac{1}{\gamma_{t, t}} \gamma_t \otimes \mathbf{X} \quad (16)$$

respectively, where  $\mathbf{1}_t$  is the unit vector with the  $t$ th entry set equal to one, and  $\gamma_t$  is the  $t$ th column of  $\mathbf{\Gamma}$ .

### III. RESULTS

#### A. Rank and singular value spectrum of the dynamic lead field mapping $\mathbf{D}_t$

We used the dynamic lead field mapping construct in conjunction with an MRI-based MEG forward model from a human subject (see Section V-C and V-D in **Supplemental Materials**) to estimate the number of independent sources we can recover in a model that includes local spatiotemporal cortical dynamics (Eq. 1). In order to avoid the computational challenges associated with finding the rank and singular values of the very large  $\mathbf{D}_t$  matrix (Eq. 11), we chose to analyze the truncated versions of  $\mathbf{D}_t$  that correspond to a model that includes  $k \in [1, 2, 5, 10, 20]$  measurements into the past and future of a given time  $t$ ,  $[\mathbf{y}_{t-k}, \dots, \mathbf{y}_t, \dots, \mathbf{y}_{t+k}]$ . The resulting mappings, which we denote by  $\mathbf{D}_t(k)$ , are given by:

$$\mathbf{D}_t(k) = [\mathbf{F}_b^{k'} \mathbf{X}', \dots, \mathbf{F}_b' \mathbf{X}', \mathbf{X}', \mathbf{F}' \mathbf{X}', \dots, \mathbf{F}^{k'} \mathbf{X}']'. \quad (17)$$

This arrangement was chosen to compare the relative contributions of increasing numbers of observations to the number of independent sources that can be effectively recovered. Furthermore, since the matrices  $\mathbf{D}_t(k)$  correspond to a truncated version of the much larger matrix  $\mathbf{D}_t$ , our computations serve as a lower bound on the number of sources that could

be recovered from the time series of measurements in the complete interval  $[1, 2, \dots, T]$ .

Figure 2 (left) and Table I show the evolution of the singular value spectrum and the rank of the  $\mathbf{D}_t(k)$  matrix, respectively, as the number of observed data points  $2k + 1$  increases. With each successive increase in  $k$ , the singular value spectrum and the rank of the  $\mathbf{D}_t(k)$  increases monotonically, reaching a value of  $\text{rank}(\mathbf{D}_t(20)) = 4551$ . This indicates the number of sources that could be recovered increases by up to a factor of 20 by modeling only local cortical dynamic connections.

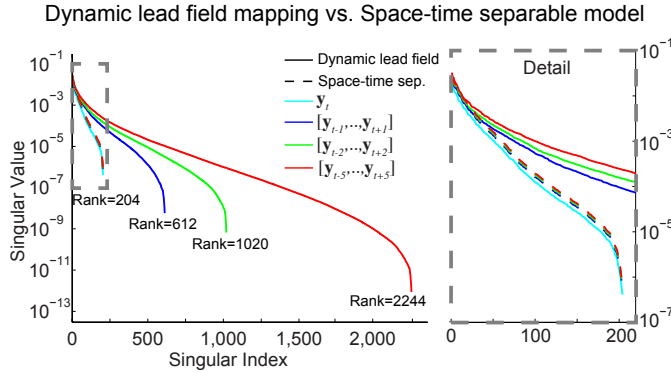


Fig. 2. **Spectrum of the dynamic lead field mapping  $\mathbf{D}_t(k)$ .** The evolution of the singular value spectra of  $\mathbf{D}_t(k)$  (solid lines) with increasing number of temporal measurements  $k$ . The spectrum of this matrix increases thus indicating an increase in the number of sources that could be recovered in the spatiotemporal dynamic model with local connections. A zoom-in is shown (right) to detail the singular value spectra of  $\mathbf{D}_t(k)^{(sts)}$  for the space-time separable (STS) source model (dashed line). In the STS source model the number of sources that could be recovered does not increase since the rank stays constant.

TABLE I  
Rank of  $\mathbf{D}_t(k)$ .

$k$	0	1	2	5	10	20
$\text{rank}(\mathbf{D}_t(k))$	204	612	1020	2244	3889	4551

Similarly, we analyzed the the number of sources that could be recovered in the *IND* and *STS* models by evaluating the rank and singular values spectrum of the  $\mathbf{D}_t^{(ind)}$  and  $\mathbf{D}_t^{(sts)}$  matrices, respectively (Eq. 16). Just as we did in the case of our spatiotemporal dynamic model, we evaluated the truncated versions of these matrices given by:

$$\begin{aligned} \mathbf{D}_t(k)^{(ind)} &= [\mathbf{0}', \dots, \mathbf{0}', \mathbf{X}', \mathbf{0}', \dots, \mathbf{0}']' \\ \mathbf{D}_t(k)^{(sts)} &= \\ &= \frac{1}{\gamma_{t,t}} [\gamma_{t-k,t} \mathbf{X}', \dots, \gamma_{t-1,t} \mathbf{X}', \mathbf{X}', \gamma_{t+1,t} \mathbf{X}', \dots, \gamma_{t+k,t} \mathbf{X}'] \end{aligned} \quad (18)$$

For both the *IND* and *STS* models, the rank of these matrices did not increase with the inclusion of more temporal measurements:  $\text{rank}(\mathbf{D}_t(k)^{(ind)}) = \text{rank}(\mathbf{D}_t(k)^{(sts)}) = 204$  for all values of  $k$ . In the case of the *STS* model, we saw that the 204 singular values of  $\mathbf{D}_t(k)^{(sts)}$  were slightly increased (Fig. 2, right). However, such increase was uniform across singular values, as it can be shown that the vector with the ordered singular values of  $\mathbf{D}_t(k)^{(sts)}$  is equal to

a scaled version of the vector with the singular values of  $\mathbf{X}$ :  $\text{svd}(\mathbf{D}_t(k)^{(sts)}) \propto \text{svd}(\mathbf{X})$ . Importantly, the fact that the rank of  $\mathbf{D}_t(k)^{(ind)}$  and  $\mathbf{D}_t(k)^{(sts)}$  does not increase can be seen from the Equation (18). In both cases, these matrices are obtained by stacking the lead field matrix  $\mathbf{X}$  above and below with either a matrix of zeros or scaled versions of  $\mathbf{X}$ .

### B. Sensitivity analysis of dynamic lead field mapping $\mathbf{D}_t$

As discussed earlier, the biophysics of the MEG/EEG forward problem dictate that some regions of the brain are more difficult to observe and measure than others: the measured signal decays with the inverse of the square distance from the sensor to the source, and in the case of MEG, sources oriented near the radial direction are magnetically silent. Figure 3 (left) illustrates this phenomenon where we show arrows representing dipole sources located in the right temporal cortex (deep dipole, white), the trough of a gyrus (mildly deep dipole, red), and the side of a gyrus (superficial dipole, yellow). In this hypothetical example, the deep white dipole located at a distance  $d_1$  from its closest sensor produces very low amplitude signals which renders it very difficult to detect. The mildly deep red dipole at a distance  $d_2$  from its closest sensor generates relatively weak signals, and this also makes it difficult to measure. In contrast to the previous two cases, the yellow dipole located on the superficial side of the gyrus at a distance  $d_3$  from its closest sensor produces strong signals which makes it easier to recover. A natural way to quantify this ease or difficulty for detecting the electric and magnetic fields at the scalp surface when only the instantaneous  $\mathbf{y}_t$  is observed, is to compute the signal power measured across sensors generated by a single active dipole source of unit amplitude. Specifically, if we fix the source vector to represent a unit amplitude active dipole at the  $i$ th cortical location, the total signal power it produces at the sensors is given by:

$$s_{i,t}(0) = \|\mathbf{X}\mathbf{1}_i\|_2, \quad (19)$$

Performing such computation for all cortical locations  $i$  results in a static profile of the *absolute* sensitivity gain of the lead field matrix  $\mathbf{X}$ . Figure 3 (right) shows this static sensitivity profile computed using lead field matrix described in the previous section. The sensitivity is overlaid on the inflated cortical surface where dark and light gray areas represent sulci and gyri, respectively. We saw that the sensitivity was highest in some portions of gyri and the more superficial cortical areas (red tones), but it was low in the troughs of sulci, insula, and inferior-frontal regions.

While the sensitivity of the lead field matrix gives us a static image of the signal quality obtained in the instantaneous measurement  $\mathbf{y}_t$ , it does not inform us of the quality of the measured signals in time when we account for the spatiotemporal dynamics of the underlying cortical activity. One way to assess such signal quality is to extend the sensitivity analysis done for the lead field matrix (Eq. 19) to the case of the dynamic lead field  $\mathbf{D}_t(k)$ . Specifically, if we assume a unit amplitude dipole is active at time  $t$  and cortical location  $i$  in our spatio-temporal dynamic model (Eq. 2), the total measured



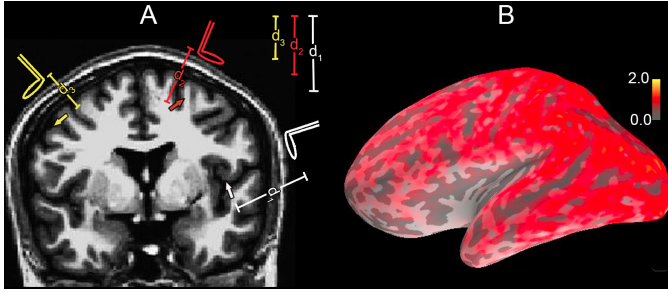


Fig. 3. **Illustration of sensitivity of the static lead field matrix to dipoles at different depths.** The white dipole located very deep in the right auditory cortex at a distance  $d_1$  from its closest sensor produces very weak signals which makes it very difficult to detectable. The red dipole near the trough of a gyrus at a distance  $d_2$  from its closest sensor produces relatively weak signals and is difficult to detect. The superficial yellow dipole in the side of a sulcus at a distance  $d_3$  from its closest sensor produces strong signals. For each dipole, the absolute sensitivity can be computed using Equation (19) resulting in the sensitivity map overlaid on the inflated cortical surface (right). The dark and light gray areas on the inflated surface represent sulci and gyri, respectively. The color-scale is multiplied by  $10^{-3}$ .

power in the time interval  $[t - k, \dots, t + k]$  across all sensor is given by:

$$s_{i,t}(k) = \|\mathbf{D}_t(k)\mathbf{1}_i\|_2, \quad (20)$$

Figure 4A (left panels) shows the absolute sensitivity of the dynamic lead field mapping as we increase the number of measurement in time  $(2k + 1)$ . We saw that the sensitivity increased in most cortical regions including sulci when more temporal information is included, but this increase was not as pronounced in deeper regions such as some portions of insula and the inferior-frontal cortex.

To quantify the improvement in the dynamic lead field sensitivity  $s_{i,t}(k)$  in relation to the static case, i.e., the static lead field matrix sensitivity  $s_{i,t}(0)$ , we computed the *relative* sensitivity gain:  $s_{i,t}(k)/s_{i,t}(0)$ . Figure 4B (right panels) shows the relative sensitivity gain of the dynamic lead field mapping. We found that the relative sensitivity increases in the majority of cortical regions with additional temporal information. Interestingly, the increments in relative sensitivity were higher in regions of low absolute sensitivity, such as sulci, insula, and inferior frontal cortex (yellow tones).

We performed an equivalent sensitivity analysis assuming that the cortical activation was generated by the space-time separable (STS) model. To compute this sensitivity, we used Equation (20) but in this case replaced  $\mathbf{D}_t(k)$  with the truncated matrix  $\mathbf{D}_t(k)^{(sts)}$  (Eq. 18). Figure 5A (left panels) shows the absolute sensitivity obtained in the STS model. We saw an increase in sensitivity in most cortical regions, but this increase did not appear as broadly spread as it was in the case of the dynamic lead field mapping (Fig. 4A). Figure 5B shows the relative sensitivity in the (STS) model. The relative sensitivity in the STS model increased with the number of temporal measurements, but this increase was spatially uniform across cortex. Compared to the relative sensitivity of the dynamic lead field (Fig. 4B), the increments in relative sensitivity under the STS model were not as pronounced in regions that are more

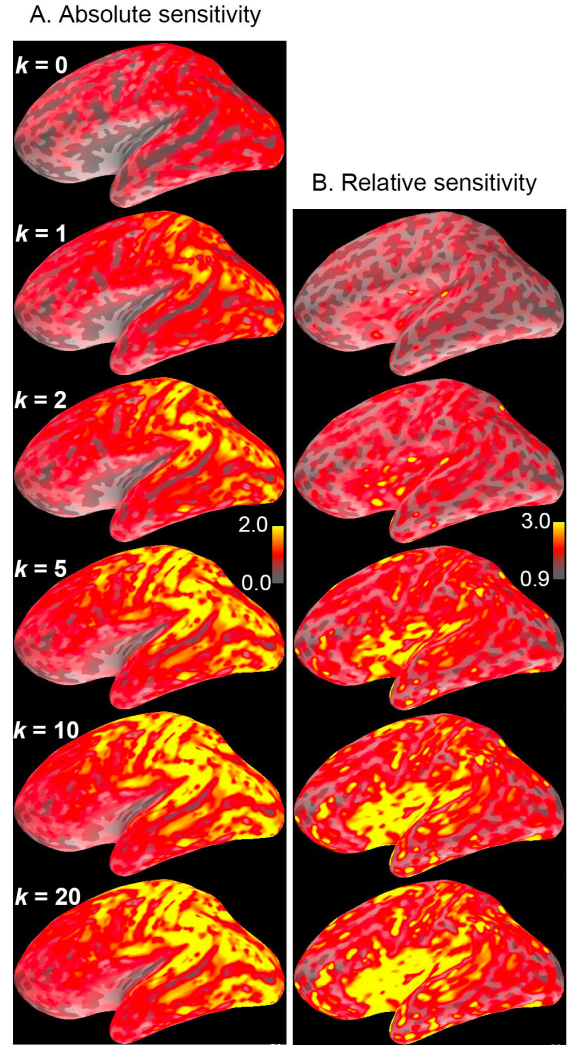


Fig. 4. **Sensitivity analysis of the dynamic lead field mapping.** A: Absolute sensitivity of  $\mathbf{D}_t(k)$  as a function of the number of incorporated measurements  $2k + 1$ . When  $k = 0$ , the sensitivity of the static lead field matrix  $\mathbf{X} = \mathbf{D}_t(0)$  is shown. For  $k \geq 1$ , the sensitivity increases in most cortical regions including sulci. The color-scale is multiplied by  $10^{-3}$ . B: The relative sensitivity of the dynamic lead field mapping increases with  $k$  in most cortical regions. This relative increase is more pronounced in areas with low static sensitivity such as sulci, insula, and inferior frontal cortex.

difficult to detect from a static point of view, such as sulci, insula, and inferior-frontal regions.

Figure 6 shows the difference in absolute sensitivity between our spatiotemporal dynamic model and the STS model for each value of  $k$ :  $s_{i,t}(k) - s_{i,t}(k)^{(sts)}$ . We should note that the color scale in this case is different from that in panel A of Figures 4 and 5: the red and blue tones indicate small positive and negative differences of  $\sim 0.2 \times 10^{-3}$ , respectively; large positive difference of  $\sim 2 \times 10^{-3}$  are shown in yellow, while large negative differences are shown in light blue. For the majority of cortical areas, and all values of  $k$ , the sensitivity was higher in the spatiotemporal dynamic model as indicated by the red and yellow tones. The small areas where the sensitivity of the STS model was slightly higher (blue areas where the difference is  $\sim -0.2 \times 10^{-3}$ ) were mainly represented by superficial portions of sulci which are indeed already

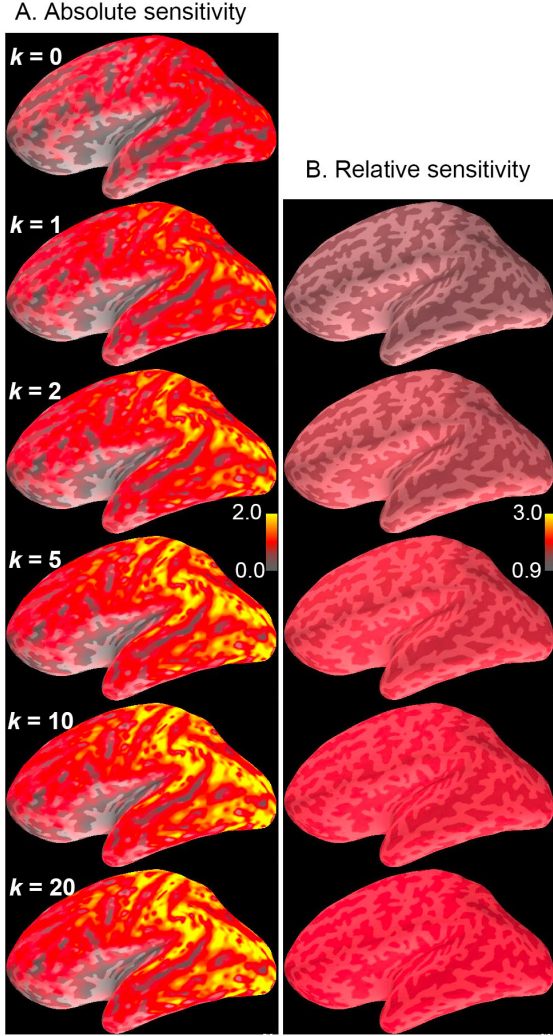


Fig. 5. **Sensitivity analysis of the space-time separable model.** A: Absolute sensitivity for the space-time separable model  $\mathbf{D}_t(k)^{(sts)}$  as a function of  $k$ . The sensitivity increases in this case but is not as widespread nor as high as it was in the case of the dynamic lead field (Fig. 4A). The color-scale is multiplied by  $10^{-3}$ . B: The increments in relative sensitivity for the space-time separable model are spatially uniform and not as pronounced in regions that are difficult to detect such as sulci, insula, and inferior frontal cortex.

detectable from a static point of view, i.e., if we only consider the immediate measurement for analysis. Interestingly, for values of  $k \geq 5$  the sensitivity in the spatiotemporal dynamic model is higher in deeper areas such as sulci and insular cortex, which are precisely the areas that are more difficult to detect from a static point of view.

We saw in the previous section that, in a model that includes local spatiotemporal cortical dynamics, adding temporal measurements ( $k$ ) increases the rank of the  $\mathbf{D}_t(k)$  matrices. We therefore wanted to characterize the dynamic lead field sensitivity that corresponded to the rank increments, which thus represent the sensitivity gains in the newly accessible dimensions. Specifically, if we consider two dynamic lead field matrices  $\mathbf{D}_t(k_1)$  and  $\mathbf{D}_t(k_2)$  with  $k_1 < k_2$ , the sensitivity in the dimension accessible at  $k_2$  but not at  $k_1$  can be obtained by: first projecting the rows of  $\mathbf{D}_t(k_2)$  onto the null space of  $\mathbf{D}_t(k_1)$ ; and then compute the sensitivity of the matrix

Dynamic lead field vs. space-time separable.

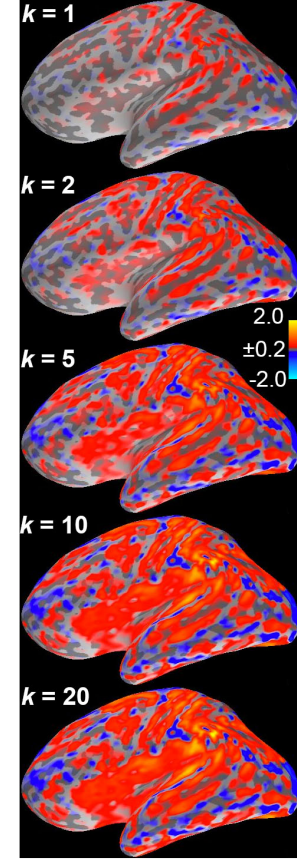


Fig. 6. **Difference in absolute sensitivity between dynamic lead field and space-time separable models.** For the majority of cortical areas the sensitivity in the dynamic lead field mapping is higher than that of the STS model (red and yellow tones). For  $k \geq 5$ , the difference in sensitivity is more pronounced in favor of the dynamic lead field mapping in deeper cortical areas. The color-scale is multiplied by  $10^{-3}$ .

constructed with the projected rows. Figure 7A shows the sensitivity in the newly accessible areas for the pairs of values corresponding to consecutive number in  $k \in [1, 2, 5, 10, 20]$ . As  $k$  increased, the distribution of sensitivity for the newly accessible dimensions was highest in some portions of sulci, insula, and inferior-frontal regions, which are poorly detected areas in the static case (Fig. 7A). We performed an equivalent analysis for the STS model. In this case there were no gains in sensitivity in the previously inaccessible dimensions (Fig. 7B). This results from the fact that the rank of the  $\mathbf{D}_t(k)^{(sts)}$  matrix does not increase irrespective to the number of analyzed measurements ( $k$ ).

#### IV. CONCLUSIONS AND DISCUSSION

In this work we have shown that the number of independent sources that can be recovered from MEG recordings can increase up to a factor of  $\sim 20$  by modeling linear source dynamics and local cortical connections. To do this, we developed and analyzed the concept of the *dynamic lead field mapping*. This dynamic mapping extends the static observation model that originates from the lead field biophysics



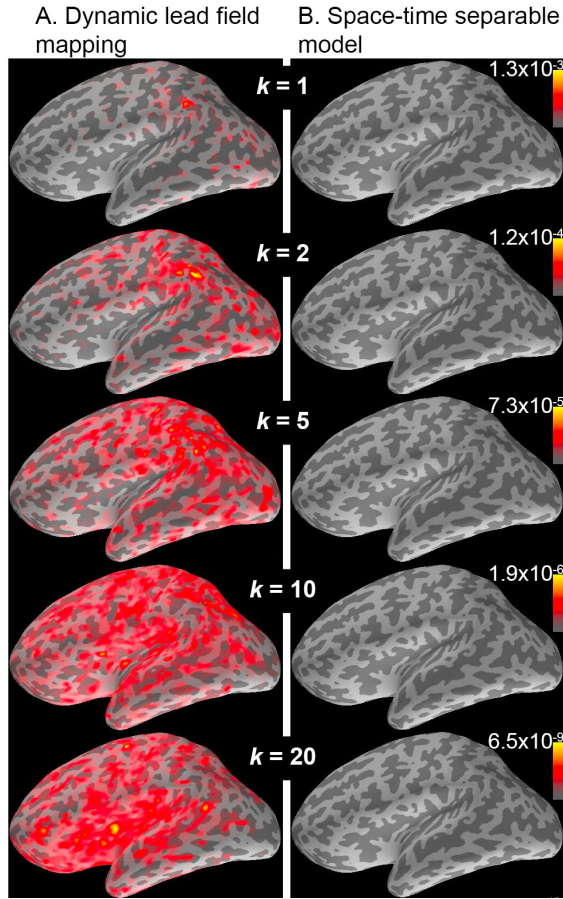


Fig. 7. **Sensitivity in newly accessible dimensions: Dynamic lead field vs. space-time separable model.** A: The sensitivity of the dynamic lead field mapping in the newly accessible dimension increases as a function of  $k$  and is highest in some portions of sulci, insula, and inferior frontal regions. B: In the case of the space-time separable model, there are no sensitivity gains in previously inaccessible dimensions since the rank of the matrix  $\mathbf{D}_t(k)^{(sts)}$  does not increase.

(Eq. 3) to account for spatiotemporal dynamics [16]–[22]. The dynamic lead field mapping expresses the relation between the cortical source vector at any given time  $t$  and the MEG/EEG measurements in the complete experiment interval  $[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$ , creating an observation model for the complete time series of measurements that maintains the crucial signal-plus-noise structure present in its static counterpart. Since the rank of the lead field matrix determines the number of independent variables we can expect to recover from an individual measurement in time, the rank of the dynamic lead field mapping determines the number of variables we can recover in the dynamic case from the complete time series of observations. Therefore, in a typical MRI-based source model with  $\sim 5124$  dipole sources and  $\sim 204$  sensors, while the static (IND) or space-time separable (STS) models can recover at most  $\sim 4\%$  ( $204/5124 \times 100$ ) of the independent variables, the approximate  $4551/204 \approx 20$  fold rank improvement of the dynamic lead field implies that we could potentially recover up to  $\sim 89\%$  ( $4551/5124 \times 100$ ) of the independent variables if we were to model the brain’s source activity as a spatiotemporal dynamic process.

While the static lead field matrix provides information about

the source vector from a single measurement at a given point in time, the dynamic lead field mapping provides additional information from past and future measurements that reflect the trajectory of the dynamic cortical state. Along this trajectory, activity hidden within a “blind spot” of the static lead field at a given time can evolve into a “visible” portion of the static lead field due to the brain’s high level of connectivity. Since the dynamic lead field mapping explicitly describes how the measurement at one time point contains information about the cortical state at another time point, this mapping effectively shows how information can flow from blind spots in the static lead field to more visible locations. In principle, depending on the cortical connectivity and dynamics, it is then possible to access and potentially estimate a much larger number of independent variables. Our work shows that modeling the simplest spatiotemporal relationships could dramatically improve source localization. Moreover, it suggests that more comprehensive models of connectivity and spatiotemporal dynamics could provide even greater improvements.

We performed a sensitivity analysis of the dynamic lead field mapping to determine how sensitivity in different cortical areas changes as we increased the number of measurements in time. This analysis showed that, in our spatiotemporal dynamic paradigm, brain areas that are difficult to detect from a static point of view, such as those with low sensitivity in troughs of sulci, insula, and inferior-frontal regions—“blind spots” as described above—become more accessible with the addition of temporal measurements. Moreover, we showed that this increase in sensitivity in deep sources does not occur in static (IND) or space-time separable (STS) models, suggesting that activity within “blind spots” become accessible only through their connection to more visible regions. This could explain the improvements in MEG/EEG source localization accuracy obtained via algorithms that explicitly model source dynamics with local cortical connections [11], [13].

We envision a possible way in which our dynamic lead field construct could be extended to more general source models, to potentially include non-Gaussian models or more realistic spatiotemporal neural source dynamics. In principle, this could be achieved by taking techniques from control theory designed to analyze deterministic systems, and adapting them to characterize properties of stochastic dynamical systems. In control theory, a basic question is to determine if different initial states of a deterministic dynamical system necessarily produce different measurements. When this is the case, the system is called *observable*, since the system’s initial states can be distinguished based on measurements alone [24]. In our linear dynamic state-space model (Eqs. 2 and 3), for example, assuming that the state input  $\omega_t$  is a priori known and deterministic, and that the measurements are noiseless ( $\varepsilon_t = 0$ ), effectively converts our model into a deterministic system. In this deterministic case, an unknown initial state  $\beta_0$  of the system can be distinguished based on noiseless measurements when the rank of the so-called observability matrix  $\mathcal{O} = [\mathbf{X}', \mathbf{F}'\mathbf{X}', \dots, \mathbf{F}^{p-1'}\mathbf{X}']'$  is equal to the dimension of the state ( $p$ ). We can see here that the observability matrix  $\mathcal{O}$  is related to the mapping given by the dynamic lead field from the state  $\beta_t$  to only the present and future observations

$\{y_t, y_{t+1}, \dots, y_{t+p-1}\}$  (this can be seen in Equation (11)). We therefore hypothesize that the dynamic lead field concept developed here could be extended by incorporating insights from the analysis of observability in deterministic nonlinear systems [25], and by applying the time-reversed representation of nonlinear stochastic dynamic systems [26].

A long-standing conjecture in functional neuroimaging is that the integration of different imaging modalities, such as MEG, EEG, fMRI, and DOT, could improve spatiotemporal resolution, due to the complementary nature of the physics and physiology underlying these modalities. Extensions of our analysis to multimodal neuroimaging could provide a formal paradigm to characterize and maximize the spatiotemporal resolution that can be obtained from multimodal data. For example, an immediate multimodal effort could be aimed at including long-distance connections—both between distant cortical locations, and between cortex and low sensitivity subcortical areas—by using anatomical connectivity derived from diffusion MRI [20]. We predict that the inclusion of these long-distance connections in a stochastic dynamics framework could have a dramatic impact on the number and location of source parameters that could be recovered from MEG/EEG time series data. While the focus of our work is in the field of neuroimaging, our dynamic lead field approach could have application in other areas where an underlying spatiotemporal stochastic dynamical system is observed through noisy serial measurements. Therefore, we hypothesize that our approach could find applications in areas such as geophysics, network theory, and epidemiology, where physical spatiotemporal dynamic relationships may exist, but are not yet exploited in inverse solutions.

Our analysis and results have been focused on the number, spatial distribution, and sensitivity of sources that can be recovered under a spatiotemporal dynamic framework. In practical applications, model misspecification, as well as the signal-to-noise of the measurements, will likely limit performance gains to some extent below the levels we have reported. However, the large improvements in rank and sensitivity that we have observed, even under a very simple spatiotemporal model, suggest that substantial performance improvements would be achievable in practical source localization applications. In future work, we will analyze source localization performance under this spatiotemporal framework and characterize the influence of different spatiotemporal model choices.

#### ACKNOWLEDGMENT

We would like to thank Dr. Simona Temereanca for providing the MEG and MRI data used in this paper.

#### REFERENCES

- [1] M. S. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, "Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain," *Rev. Modern Phys.*, vol. 65, no. 2, pp. 413–497, 1993.
- [2] M. S. Hämäläinen and R. J. Ilmoniemi, "Interpreting magnetic fields of the brain: Minimum norm estimates," *Med. Biol. Eng. Comput.*, vol. 32, no. 1, pp. 35–42, 1994.
- [3] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain," *Int. J. Psychophysiol.*, vol. 18, no. 1, pp. 49–65, Oct. 1994.
- [4] A. Nummenmaa, T. Auranen, M. S. Hämäläinen, I. P. Jääskeläinen, M. Sams, A. Vehtari, and J. Lampinen, "Automatic relevance determination based hierarchical Bayesian MEG inversion in practice," *NeuroImage*, vol. 37, no. 3, pp. 876–889, Sep. 2007.
- [5] D. P. Wipf and S. S. Nagarajan, "A unified Bayesian framework for MEG/EEG source imaging," *NeuroImage*, vol. 44, no. 3, pp. 947–966, Feb. 2009.
- [6] S. Baillet and L. Garnero, "A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 5, pp. 374–385, 1997.
- [7] K. J. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. J. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout, "Multiple sparse priors for the M/EEG inverse problem," *NeuroImage*, vol. 39, no. 3, pp. 1104–1120, Feb. 2008.
- [8] W. Ou, M. S. Hämäläinen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, no. 3, pp. 932–946, Feb. 2009.
- [9] A. Bolstad, B. D. Van Veen, and R. Nowak, "Space-time event sparse penalization for magneto-/electroencephalography," *NeuroImage*, vol. 46, no. 4, pp. 1066–1081, Jul. 2009.
- [10] A. Gramfort, D. Strohmeier, J. Haueisen, M. S. Hämäläinen, and M. Kowalski, "Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations," *NeuroImage*, vol. 70, pp. 410–422, Apr. 2013.
- [11] A. Galka, O. Yamashita, T. Ozaki, R. Biscay, and P. A. Valdes-Sosa, "A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering," *NeuroImage*, vol. 23, no. 2, pp. 435–453, 2004.
- [12] C. J. Long, P. L. Purdon, S. Temereanca, N. U. Desai, M. S. Hämäläinen, and E. N. Brown, "State-space solutions to the dynamic magnetoencephalography inverse problem using high performance computing," *Ann. Appl. Stat.*, vol. 5, no. 2B, pp. 1207–1228, Jun. 2011.
- [13] C. Lamus, M. S. Hämäläinen, S. Temereanca, E. N. Brown, and P. L. Purdon, "A spatiotemporal dynamic distributed solution to the MEG inverse problem," *NeuroImage*, vol. 63, no. 2, pp. 894–909, 2012.
- [14] A. M. Dale and M. I. Sereno, "Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach," *J. Cogn. Neurosci.*, vol. 5, no. 2, pp. 162–176, 1993.
- [15] S. P. Ahlfors, J. Han, J. W. Belliveau, and M. S. Hämäläinen, "Sensitivity of MEG and EEG to source orientation," *Brain Topogr.*, vol. 23, no. 3, pp. 227–232, Jul. 2010.
- [16] T. H. Bullock, M. C. McClune, J. Z. Achimowicz, V. J. Iragui-Madoz, R. B. Duckrow, and S. S. Spencer, "Temporal fluctuations in coherence of brain waves," *PNAS*, vol. 92, no. 25, pp. 11 568–11 572, Dec. 1995.
- [17] A. Destexhe, D. Contreras, and M. Steriade, "Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states," *J. Neurosci.*, vol. 19, no. 11, pp. 4595–4608, Jun. 1999.
- [18] D. A. Leopold, Y. Murayama, and N. K. Logothetis, "Very slow activity fluctuations in monkey visual cortex: Implications for functional brain imaging," *Cereb. Cortex*, vol. 13, no. 4, pp. 422–433, Apr. 2003.
- [19] P. L. Nunez, *Neocortical dynamics and human EEG rhythms*, 1st ed. New York: Oxford University Press, 1995.
- [20] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the structural core of human cerebral cortex," *PLoS Biol.*, vol. 6, no. 7, pp. 1479–1493, 2008.
- [21] C. J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J. P. Thiran, R. Meuli, and P. Hagmann, "Predicting human resting-state functional connectivity from structural connectivity," *PNAS*, vol. 106, no. 6, pp. 2035–2040, Feb. 2009.
- [22] M. P. van den Heuvel, R. C. W. Mandl, R. S. Kahn, and H. E. Hulshoff Pol, "Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain," *Hum. Brain Mapp.*, vol. 30, no. 10, pp. 3127–3141, Oct. 2009.
- [23] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*, 1st ed. New York: Springer, Aug. 2013.
- [24] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*, 1st ed. Upper Saddle River: Prentice Hall, 2000.
- [25] E. D. Sontag, *Mathematical control theory*, 2nd ed., ser. Deterministic finite dimensional systems. New York: Springer, Jul. 1998.
- [26] B. D. O. Anderson, "Reverse-time diffusion equation models," *Stoch. Proc. Appl.*, vol. 12, no. 3, pp. 313–326, May 1982.

## V. Supplemental Materials

### An Analysis of How Spatiotemporal Dynamic Models of Brain Activity Could Improve MEG/EEG Inverse Solutions

Camilo Lamus, Matti S. Hämäläinen, Emery N. Brown, and Patrick L. Purdon

#### A. Steady-state covariance of the source process $\beta_t$

In this section we obtain the steady-state source covariance  $\mathbf{C} = \text{Cov}(\beta_t)$  of our dynamic source models (Eq. 2), and show that this covariance matrix is invertible. The matrix  $\mathbf{F}$  is stable, i.e., the magnitude of its largest eigenvalue is equal to  $\phi < 1$ , since  $\mathbf{F}/\phi$  is a stochastic matrix (it has nonnegative entries with the sum of its rows equal to 1) and because the largest eigenvalue of a stochastic matrix is 1. Since  $\mathbf{F}$  is stable, the process covariance reaches a steady-state,  $\mathbf{C} = \lim_{t \rightarrow \infty} \text{Cov}(\beta_t)$ , which is given by the solution of the Lyapunov equation  $\mathbf{C} = \mathbf{F}\mathbf{C}\mathbf{F}' + \mathbf{Q}$ . Furthermore, since  $\mathbf{Q}$  is positive definite and invertible, so it is the source covariance  $\mathbf{C}$  [24]. Therefore, we can assume that the process  $\beta_t$  has reached a steady-state where its spatial covariance matrix  $\mathbf{C} = \text{Cov}(\beta_t)$  is time invariant and invertible.

We should note that in our analysis we can parametrized the covariance matrix  $\mathbf{Q} = [\lambda \text{tr}(\hat{\Sigma})/n]^{-1} \text{diag}(\nu_1, \nu_2, \dots, \nu_p)$ , where  $\hat{\Sigma} = \mathbf{X}'\mathbf{X}/n$  is the sample covariance of the rows of  $\mathbf{X}$ ,  $\lambda > 0$ , and  $\nu_i > 0$ . Doing so would allow us to interpret  $\lambda$  as the inverse of the power signal-to-noise ratio ( $\text{SNR}^2$ ). This is because if we assume the matrix  $\mathbf{F} = \phi\mathbf{I}$ , and  $\nu_i = (1 - \phi^2)$  ( $i = 1, \dots, p$ ), the steady state covariance of  $\beta_t$  becomes  $\mathbf{C} = [\lambda \text{tr}(\hat{\Sigma})/n]^{-1} \mathbf{I}$ . If we define the power signal-to-noise ratio as  $\text{SNR}^2 = \mathbb{E}[\|\mathbf{X}\beta_t\|^2]/\mathbb{E}[\|\varepsilon_t\|^2]$ , then  $\text{SNR}^2 = \text{tr}(\mathbf{X}'\mathbf{X})[\lambda \text{tr}(\hat{\Sigma})/n]^{-1}/n = 1/\lambda$ .

#### B. The dynamic lead field mapping projects $[\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_T]'$ onto $\beta_t$

In this section we show that the observation model for the complete set of measurements  $[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$  under the dynamic lead field mapping possesses the desired *signal plus noise* structure in which the source vector at time  $t$  is independent of the *noise* term (Eq. 10). In order to do this, we will show that the matrices  $\mathbf{P}_{t \pm k, t}$  that make up the blocks of the dynamic lead field mapping are indeed matrices projecting the corresponding measurement  $\mathbf{y}_{t \pm k}$  onto the source vector  $\beta_t$ . Doing so would imply that the dynamic lead field mapping achieves the desired independence condition.

We should first point to a few facts about the linear projection  $\mathbf{P}_{t \pm k, t} \beta_t$  of the measurement vector  $\mathbf{y}_{t \pm k}$  onto the source vector  $\beta_t$ . By definition, the projection error  $\mathbf{e}_{t \pm k, t} = \mathbf{y}_{t \pm k} - \mathbf{P}_{t \pm k, t} \beta_t$  is uncorrelated with  $\beta_t$  [24]:

$$\mathbb{E}[(\mathbf{y}_{t \pm k} - \mathbf{P}_{t \pm k, t} \beta_t) \beta_t'] = \mathbf{0}. \quad (\text{S1})$$

In addition, since  $\mathbf{y}_{t \pm k}$  and  $\beta_t$  are jointly Gaussian, the projection error  $\mathbf{e}_{t \pm k, t}$  and  $\beta_t$  are also jointly Gaussian. Therefore, the projection error and the source vector are uncorrelated and jointly Gaussian, and thus are independent. As a result, we can use the definition of the projection error to express any measurement vector  $\mathbf{y}_{t \pm k}$  as the sum of the projection  $\mathbf{P}_{t \pm k, t} \beta_t$  and projection error  $\mathbf{e}_{t \pm k, t}$ :

$$\mathbf{y}_{t \pm k} = \mathbf{P}_{t \pm k, t} \beta_t + \mathbf{e}_{t \pm k, t}.$$

If we do this for all values of  $k$  we obtain an observation model for the complete time series of measurements where the source vector  $\beta_t$  is indeed independent of the noise term  $[\mathbf{e}'_{1,t}, \mathbf{e}'_{2,t}, \dots, \mathbf{e}'_{T,t}]'$ :

$$\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{t-1} \\ \mathbf{y}_t \\ \mathbf{y}_{t+1} \\ \vdots \\ \mathbf{y}_T \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{P}_{1,t} \\ \vdots \\ \mathbf{P}_{t-1,t} \\ \mathbf{P}_{t,t} \\ \mathbf{P}_{t+1,t} \\ \vdots \\ \mathbf{P}_{T,t} \end{bmatrix}}_{\text{Signal}} \beta_t + \underbrace{\begin{bmatrix} \mathbf{e}_{1,t} \\ \vdots \\ \mathbf{e}_{t-1,t} \\ \mathbf{e}_{t,t} \\ \mathbf{e}_{t+1,t} \\ \vdots \\ \mathbf{e}_{T,t} \end{bmatrix}}_{\text{Noise}}.$$

Therefore, if we show that the choices we made for the  $\mathbf{P}_{t \pm k, t}$  matrices for the dynamic lead field mapping in Section II-B, namely,

$$\begin{aligned} \mathbf{P}_{t,t} &= \mathbf{X}, \\ \mathbf{P}_{t+k,t} &= \mathbf{X}\mathbf{F}^k, \\ \mathbf{P}_{t-k,t} &= \mathbf{X}\mathbf{F}_b^k, \end{aligned}$$

satisfy the orthogonality condition (Eq. S1), it would follow that the *signal* and *noise* portions in Equation (10) are independent.

In the case where the projection matrix corresponds to the present observation ( $k = 0$ ), we substitute Equation (3) in Equation (S1), and note that setting  $\mathbf{P}_{t,t} = \mathbf{X}$  achieves the desired orthogonality condition since the source vector  $\beta_t$  is independent of the measurement noise  $\varepsilon_t$ :

$$\begin{aligned} \mathbb{E}[(\mathbf{y}_t - \mathbf{P}_{t,t}\beta_t)\beta_t'] &= \mathbb{E}[(\mathbf{X}\beta_t + \varepsilon_t - \mathbf{P}_{t,t}\beta_t)\beta_t'] \\ &= \mathbb{E}[\varepsilon_t\beta_t'] \\ &= \mathbf{0}. \end{aligned}$$

To analyze the case where the measurement corresponds to a future observation, we should first make note of two facts in relation to the source vector process. The first is that, by definition, the input process to the source dynamics is independent across time, i.e.,  $\omega_t$  is independent of  $\omega_{t \pm j}$  for  $j \in [1, 2, \dots]$ . The second is that given the recursive definition of the source process (Eq. 2), the source vector at a particular time  $t$  is a function of only the present input  $\omega_t$  and past inputs  $\omega_{t-j}$ ,  $j \in [1, 2, \dots]$ . Because of these two facts, the source vector at time  $t$ ,  $\beta_t$ , is independent of the future inputs  $\omega_{t+j}$ ,  $j \in [1, 2, \dots]$ .

With this result at hand, we proceed to deduce the future projection matrices  $\mathbf{P}_{t+k,t}$ ,  $k \in [1, 2, \dots]$  by using the forward iteration in Equation (5):  $\beta_{t+k} = \mathbf{F}^k\beta_t + \sum_{j=1}^k \mathbf{F}^{k-j}\omega_{t+j}$ . We substitute Equations (3) in Equation (S1), then substitute Equation (5) in the obtained result, and note that by setting  $\mathbf{P}_{t+k,t} = \mathbf{X}\mathbf{F}^k$  we achieve the orthogonality condition since  $\beta_t$  is independent of both  $\varepsilon_{t+k}$ , and  $\omega_{t+j}$ ,  $j \in [1, 2, \dots]$ :

$$\begin{aligned} \mathbb{E}[(\mathbf{y}_{t+k} - \mathbf{P}_{t+k,t}\beta_t)\beta_t'] &= \mathbb{E}[(\mathbf{X}\beta_{t+k} + \varepsilon_{t+k} - \mathbf{P}_{t+k,t}\beta_t)\beta_t'] \\ &= \mathbb{E}[(\mathbf{X}\mathbf{F}^k\beta_t + \sum_{j=1}^k \mathbf{X}\mathbf{F}^{k-j}\omega_{t+j} + \varepsilon_{t+k} - \mathbf{P}_{t+k,t}\beta_t)\beta_t'] \\ &= \mathbb{E}[(\sum_{j=1}^k \mathbf{X}\mathbf{F}^{k-j}\omega_{t+j} + \varepsilon_{t+k})\beta_t'] \\ &= \mathbf{0}. \end{aligned}$$

To analyze the past observations  $\mathbf{y}_{t-k}$ ,  $k \in [1, 2, \dots]$ , we proceed similarly to the case of the future projections but use instead the *backwards Markovian* representation (Eq. 7) and its  $k$ -step past iteration (Eq. 8)<sup>2</sup>:  $\beta_{t-k} = \mathbf{F}_b^k\beta_t + \sum_{j=1}^k \mathbf{F}_b^{k-j}\omega_{t-j}^b$ . We should note that in the *backwards Markovian* representation (Eq. 7), the source vector  $\beta_t$  is a function of only the present  $\omega_t^b$  and future inputs  $\omega_{t+j}^b$ ,  $j \in [1, 2, \dots]$ . In addition, given that the backwards input process  $\omega_t^b$  is independent across time, in this equivalent representation the source vector  $\beta_t$  is independent of the past inputs  $\omega_{t-j}^b$ ,  $j \in [1, 2, \dots]$ . Using this result, we can readily obtain the past projection matrices. We use Equations (3), (S1), and (8) to deduce that by setting  $\mathbf{P}_{t-k,t} = \mathbf{X}\mathbf{F}_b^k$  we obtain the desired orthogonality condition since  $\beta_t$  is independent of both  $\varepsilon_{t-k}$ , and  $\omega_{t-j}^b$ ,  $j \in [1, 2, \dots]$ :

$$\begin{aligned} \mathbb{E}[(\mathbf{y}_{t-k} - \mathbf{P}_{t-k,t}\beta_t)\beta_t'] &= \mathbb{E}[(\mathbf{X}\beta_{t-k} + \varepsilon_{t-k} - \mathbf{P}_{t-k,t}\beta_t)\beta_t'] \\ &= \mathbb{E}[(\mathbf{X}\mathbf{F}_b^k\beta_t + \sum_{j=1}^k \mathbf{X}\mathbf{F}_b^{k-j}\omega_{t-j}^b + \varepsilon_{t-k} - \mathbf{P}_{t-k,t}\beta_t)\beta_t'] \\ &= \mathbb{E}[(\sum_{j=1}^k \mathbf{X}\mathbf{F}_b^{k-j}\omega_{t-j}^b + \varepsilon_{t-k})\beta_t'] \\ &= \mathbf{0}. \end{aligned}$$

### C. Ethics statement

Human studies were approved by the Human Research Committee of the Massachusetts General Hospital, Boston, MA.

### D. Data description and preprocessing

In our analysis, the lead field matrix  $\mathbf{X}$  was computed with the *MNE* software using an MRI-based boundary element model from a human subject [1], [14], with  $p = 5124$  dipole sources oriented normal to the cortical surface. This arrangement yielded an average distance between nearest neighbors of  $\sim 6.2$  mm, resulting in a model that is consistent with intracranial

<sup>2</sup>We should point out that directly reversing the time direction of the equation defining the source spatiotemporal dynamics (Eq. 2), i.e.,  $\beta_{t-1} = \mathbf{F}^{-1}\beta_t - \mathbf{F}^{-1}\omega_t$ , instead of using the *backwards Markovian* model does not yield the desired orthogonality condition. This is because the  $k$ -step backwards iteration resulting from this direct time reversal, namely,  $\beta_{t-k} = \mathbf{F}^{-k}\beta_t - \sum_{j=0}^{k-1} \mathbf{F}^{j-k}\omega_{t-j}$ , contains the terms  $\omega_t$  and  $\omega_{t-j}$  which are not independent of the source vector  $\beta_t$ .



electrophysiology and neuroanatomy [16]–[19]. Due to the difference in orders of magnitude and physical units in the elements of the lead field matrix  $\mathbf{X}$  corresponding to planar gradiometers and magnetometers, we restricted our analysis to the 204 planar gradiometers of the Neuromag Vectorview system at the Massachusetts General Hospital. The  $\mathbf{F}$  matrix was specified using the nearest-neighbor dynamic formulation described by Equations (1) and (2), with the stability parameter  $\phi = 0.95$ . The state input covariance matrix  $\mathbf{Q}$  was estimated from *mu*-rhythm data from one subject using the *dMAP-EM* algorithm presented in [13]. The *mu*-rhythm originates from motor and somatosensory cortices, and consists of synchronous oscillations with 10 and 20-Hz components. Data were collected from one subject using a 306-channel Neuromag Vectorview MEG system at the Massachusetts General Hospital. The signals were acquired at 601 Hz with a bandwidth of 0.1 to 200 Hz, and downsampled to 204.8 Hz for subsequent analysis. In the *STS* model, the elements for the temporal covariance matrix  $\mathbf{\Gamma}$  were set as suggested in [7]:  $\gamma_{t \pm k, t} = \sum_{j=1}^T \exp\{-1/2[(t \pm k - j)^2 + (j - t)^2]\Delta^{-2}\psi^{-2}\}$ , where  $\Delta = 4 \times 10^{-3}$  s and  $\psi = 204.8$  Hz.